

## Final MB – Disturbing or disturbed judgement?

S. G. GOWERS, Lecturer; and J. HUBERT LACEY, Reader and Head, Adult Psychiatry Section, Academic Department of Psychiatry, Jenner Wing, St George's Hospital Medical School, London SW17 0RE

The decision concerning who is fit to practise medicine is rightly considered a serious one and universities hand this responsibility to a small number of senior clinicians. However, it is on patients that the doctor will practise and while all patients have a concept of what makes a good doctor, their views are never canvassed in the Final MB.

The aim of this study was to see if a group of patients – indeed psychiatric patients suffering from mental illness – could predict accurately, using their intuitive judgement, the final exam results. In particular, we had the notion that psychiatric patients and psychiatrists may hold similar opinions on this matter, not least because of the oft-expressed view that the emotionally disturbed and their medical attendants are difficult to distinguish!

### The study

The final clinical examination in medicine at a London medical school includes an examination in either psychiatry or paediatrics. The candidate is required either to take a history in half-an-hour and present it to the examiners, or interview a 'short case' in the presence of the examiners. The candidates also attend clinical examinations in adult medicine, surgery, and obstetrics and gynaecology.

The candidate sample comprised all those students sitting in a clinical examination in psychiatry. These 39 candidates were approximately half the total year's intake sitting their final examinations and were a random sample allocated by the medical school.

The patient group comprised a selection of 25 in-patients from two psychiatric hospitals. They were selected by psychiatric registrars as suitable for examination purposes and represented a broad range of psychiatric conditions (neurosis, functional and organic psychoses, personality disorders and alcoholism). The registrars were not aware that the present study was being done.

A written message was given to the patients by the first author prior to their meeting the candidates and which stated:

"After the examination you will be asked whether you think that the candidate should pass or fail the examination. Your views will not be passed on to either the examiners or the candidates and will not affect the candidate's result in the examination in any way".

Following the examination each patient was given a slip of paper and asked to grade the candidate on the basis of his own judgement according to the same scale used by the examiners, i.e. A = Distinction, B = Good Pass, C = Pass, D = Borderline Fail, E = Fail.

### Findings

Because the numbers of subjects were small, for the purpose of this study, Grades A and B were grouped together (good pass) for each candidate. The psychiatric patients awarded 20 candidates a good pass, 18 a bare pass and failed one candidate. The results of the patients' assessments were compared with the grades given each candidate by the medical, surgical and obstetric examiners and were subjected to  $\chi^2$ -squared analyses in pairs ( $3 \times 3$  tables) and Cohen's Kappa.

Psychiatric patients did not agree with psychiatrists about who should be a doctor, despite in many cases seeing the inter-action – often loaded – between candidate and examiner. Indeed, Table I indicates that psychiatrists do not appear to agree with anyone – patients or colleagues – concerning who should practise medicine. On the other hand, Table II shows that surgeons and obstetricians agreed closely, as perhaps befits sister specialities. Somewhat to our surprise, however, they held the

TABLE I  
A comparison between the psychiatrists' rating, the patients' assessment and the candidates' result in Medicine, Surgery, and Obstetrics & Gynaecology

	$\chi^2$	d.f.	P
Psychiatrists' grade v. patients' grade	6.8	4	n/s
Psychiatrists' grade v. physicians' grade	3.3	4	n/s
Psychiatrists' grade v. results in surgery	2.0	4	n/s
Psychiatrists' grade v. result in O & G	11.6	4	n/s

n/s =  $P > 0.02$

TABLE II  
A comparison between the psychiatric patients' assessment and the final grade in Surgery, Obstetrics & Gynaecology, and Medicine

	$\chi^2$	d.f.	P
Patients' grade v. result in surgery	39.4	4	0.0001
Patients' grade v. result in O & G	13.7	4	0.0085
Surgery v. obstetrics & gynaecology	13.8	4	0.0079
Medicine v. surgery	4.7	4	n/s
Medicine v. obstetrics & gynaecology	5.7	4	n/s

same opinion as the psychiatric patients. Only one candidate in this sample failed the examination in surgery – precisely the same candidate was failed by the psychiatric patient group, giving 100% agreement between the two groups about who should pass their finals!

When the patient group is broken down into those used for 'short' psychiatric cases and those used as a long case, a further interesting point emerges. The relationship between the psychiatric patients' assessment and the final grade in surgery exists only for the patients used as a short case ( $\chi^2 = 19.790$ ,  $P = 0.0005$ , d.f. = 4). In these cases the patient observed the interaction between examiner and candidate (in psychiatry) and while the patient's assessment did not predict the examination score in psychiatry, it was highly significantly associated with the surgical examiners' assessment of the candidates in a different examination.

These results were supported when the data were subjected to Cohen's Kappa. In particular, there is reasonable agreement between medicine and psychiatry (Kappa = 0.405; z stat for K = 2.852, significant at the 1% level) and to a lesser extent between surgery and medicine (Kappa = 0.218; z stat for K = 1.434, significant at the 10% level). It is worth noting that the psychiatrists and psychiatric patients agree quite considerably *less* than would be expected by chance! (Kappa = -0.256; z stat for K = -1.699).

### Comment

Despite the handicap of psychotic thought processes, preoccupations with anxiety or gross cognitive impairments, the psychiatric patients appear able to detect certain qualities which match very closely

the surgical and obstetric examiners' assessment of a final examination candidate. The results are evidently not based on the patients' perception of whether the candidates were correctly answering questions asked by the examiners, since the psychiatric patients' grade did not relate to that of the psychiatric examiners.

That the results are due to chance is unlikely given the levels of significance and given that the relationship is true for both surgery and for obstetrics and gynaecology, and not for medicine or psychiatry whose results are not related to those of the surgical specialities.

$\chi$  squares are more appropriate than Cohen's Kappa because, strictly speaking, inter-rater agreements are not being measured. The various clinicians and the patients are not rating the *same* aspects of the candidates' abilities. Manifestly, surgeons are rating ability in surgical subjects, obstetricians in obstetrics, and quite what the psychiatric patients are measuring is open to debate but certainly not abilities in surgery or obstetrics.

The puckish might suggest that these results could be seen to question the need for senior academics to examine medical students, if a psychiatric patient can judge results equally accurately. The authors would not take this irreverent view. However, we are all aware of how examinations disrupt the schedule of the busy clinician. This paper does suggest ways of reducing the load. Surgeons and obstetricians need not waste their time examining – their places could be filled to apparent equal effect by psychiatric patients. Psychiatrists also need not attend – no-one agrees with their views anyway! It would appear that only the physicians need join the emotionally disturbed to determine the next generation of doctors.

### Concluding remarks

This study stemmed from a sense of mischief engendered by the languor well-known to those who organise examinations. The aim was to see if the psychiatric patients, who volunteered to be examined by medical students in their Final MB, could accurately predict, using their intuitive judgement, the opinions of the examiners as to which students would make the better doctors. We found that psychiatric patients did not agree with the psychiatric examiners, thus dispelling once, and for all, the scurrilous view that, with time, psychiatrists think similarly to their patients. On the other hand, it would appear that our patients, many of whom were most disturbed, have similar opinions to our surgical colleagues and we would like to suggest that surgeon examiners could be replaced by psychiatric patients with no loss of sensitivity to the examination.